Study of the rare decays of to $B_{(s)}$ to muons pairs on Run 1 at ATLAS

Marcella Bona
(QMUL)
on behalf of the ATLAS collaboration

IoP joint annual HEPP and APP conference
March 22nd, 2016

Motivations, predictions and previous results

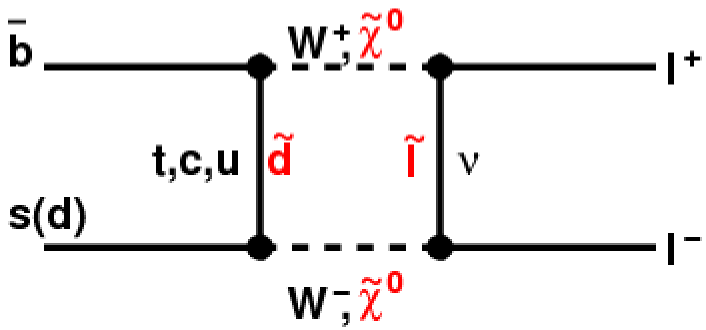
- Decays of B^0 and B^0_S into two muons have to proceed through Flavour Changing Neutral Currents (FCNC)
  $\rightarrow$ they are suppressed in the SM
- In addition, they are CKM and helicity suppressed.
- However, within the SM, they can be calculated with small theoretical uncertainties of less than 6-8%
  - latest determination includes NLO EM and NNLO QCD corrections

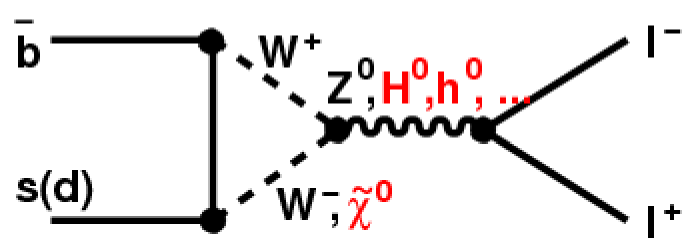
    $\mathbf{B(B^0_S \rightarrow \mu^+\mu^-) = (3.65 \pm 0.23)\ 10^{-9}}$

    $\mathbf{B(B^0 \rightarrow \mu^+\mu^-) = (1.06 \pm 0.09)\ 10^{-10}}$

Bobeth et al.,
PRL 112 (2104) 101801

- Perfect ground for indirect new physics searches:
  - virtual new physics particles can contribute to the loop
  - both enhancement and suppression effects are possible





Motivations, predictions and previous results

- Decays of B^0 and B^0_s into two muons have to proceed through Flavour Changing Neutral Currents (FCNC)
  $\rightarrow$ they are suppressed in the SM
- In addition, they are CKM and helicity suppressed.
- However, within the SM, they can be calculated with small theoretical uncertainties of less than 6-8%
  - latest determination includes NLO EM and NNLO QCD corrections

$$\mathbf{B(B^0_s \rightarrow \mu^+\mu^-) = (3.65 \pm 0.23)\ 10^{-9}}$$

$$\mathbf{B(B^0 \rightarrow \mu^+\mu^-) = (1.06 \pm 0.09)\ 10^{-10}}$$

- Perfect ground for indirect new physics searches:
  - virtual new physics particles can contribute to the loop
  - both enhancement and suppression effects are possible

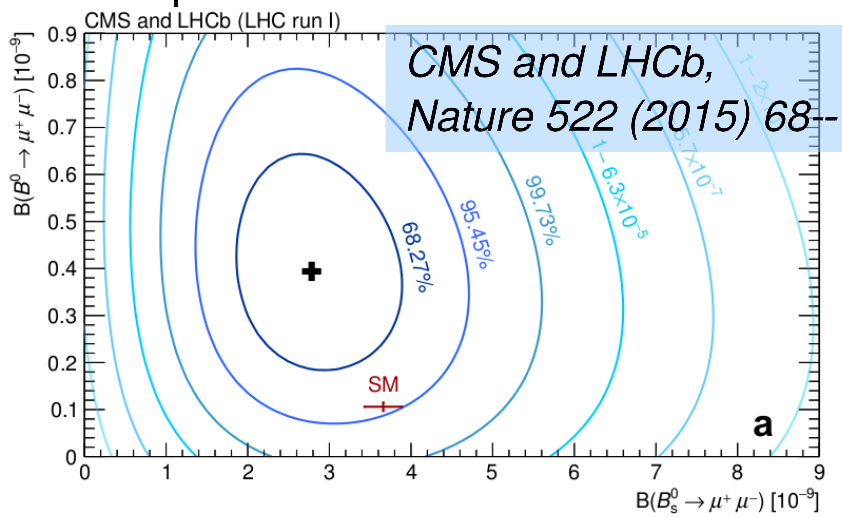
- Combination from CMS and LHCb:
  - 6σ observation for the B^0_s channel:

$$\mathbf{B(B^0_s \rightarrow \mu^+\mu^-) = (2.8^{+0.7}_{-0.6})\ 10^{-9}}$$

  - 3σ evidence for the B^0 channel:

$$\mathbf{B(B^0 \rightarrow \mu^+\mu^-) = (3.9^{+1.6}_{-1.4})\ 10^{-10}}$$

  - some tension with the SM



CMS and LHCb, Nature 522 (2015) 68--72

ATLAS analysis on full Run 1 data

- Just off the press: we present here the ATLAS study on 25 fb^{-1} of Run 1 data:
  - 4.9 fb^{-1} of 7 TeV data taken in 2011
  - 20 fb^{-1} of 8 TeV data taken in 2012
  - improved analysis strategy

- Analysis strategy:
  - dimuon triggers:
    - symmetric trigger requiring two muons with p_T > 4 GeV: good for 2011

      → prescaled in 2012, so three trigger categories

      merging asymmetric triggers *(p_T > 4 and 6 GeV)*

      and central events *(one barrel muon with $|\eta|$< 1.05)*
  - muon tracks reconstructed in both the inner detector and the muon spectrometer to improve mass resolution in muons in the end-cap region.
  - blinded analysis in the dimuon mass region: *[5166, 5526] MeV*
  - background fighting with MVA classifiers:
    - *continuum-BDT* for reducing the combinatorial background
    - *fake-BDT* for reducing the hadron misidentification as muons
  - signal extraction with a *ML fit* to the dimuon invariant mass distribution
  - normalisation with $B^{\pm} \rightarrow J/\psi K^{\pm}$ channel:

    yield, fragmentation and efficiency ratios

Analysis strategy: normalisation channel

- normalisation with $B^{\pm} \rightarrow J/\psi K^{\pm}$ channel:

$$\mathcal{B}(B^0_{(s)} \rightarrow \mu^+\mu^-) = \frac{N_{d(s)}}{\varepsilon_{\mu^+\mu^-}} \times \frac{\varepsilon_{J/\psi K^+}}{N_{J/\psi K^+}} \times \frac{f_u}{f_{d(s)}}$$

$$\times \left[\mathcal{B}(B^+ \rightarrow J/\psi K^+) \times \mathcal{B}(J/\psi \rightarrow \mu^+\mu^-)\right]$$

Analysis strategy: normalisation channel

- normalisation with $B^{\pm} \to J/\psi\, K^{\pm}$ channel:

$$\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-) = \boxed{\frac{N_{d(s)}}{\varepsilon_{\mu^+\mu^-}}} \times \boxed{\frac{\varepsilon_{J/\psi K^+}}{N_{J/\psi K^+}}} \times \frac{f_u}{f_{d(s)}}$$

$$\times \left[\mathcal{B}(B^+ \to J/\psi K^+) \times \mathcal{B}(J/\psi \to \mu^+\mu^-)\right]$$

- correction for the efficiencies of the two channels

Analysis strategy: normalisation channel

- normalisation with $B^{\pm} \rightarrow J/\psi\ K^{\pm}$ channel:

$$\mathcal{B}(B^0_{(s)} \rightarrow \mu^+\mu^-) = \frac{N_{d(s)}}{\varepsilon_{\mu^+\mu^-}} \times \frac{\varepsilon_{J/\psi K^+}}{N_{J/\psi K^+}} \times \frac{f_u}{f_{d(s)}} \times \left[\mathcal{B}(B^+ \rightarrow J/\psi K^+) \times \mathcal{B}(J/\psi \rightarrow \mu^+\mu^-)\right]$$

- correction for the efficiencies of the two channels
- correction for the different hadronisation probabilities for B^0_S and B^0 vs $B^{\pm}$
- include the $B^{\pm}$ and J/ψ branching fractions

Analysis strategy: normalisation channel

- normalisation with $B^\pm \to J/\psi\, K^\pm$ channel:

$$\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-) = \frac{N_{d(s)}}{\varepsilon_{\mu^+\mu^-}} \times \frac{\varepsilon_{J/\psi K^+}}{N_{J/\psi K^+}} \times \frac{f_u}{f_{d(s)}} \times \left[\mathcal{B}(B^+ \to J/\psi K^+) \times \mathcal{B}(J/\psi \to \mu^+\mu^-)\right]$$

- correction for the efficiencies of the two channels
- correction for the different hadronisation probabilities for B^0_S and B^0 vs $B^\pm$
- include the $B^\pm$ and J/ψ branching fractions

- Modify the above formula to take into account the three trigger categories and 2011 data
  - normalisation channel yield evaluated in each trigger and data category
  - same for the efficiency ratio

Analysis strategy: normalisation channel

- Modify the above formula to take into account the three trigger categories and 2011 data
  - normalisation channel yield evaluated in each trigger and data category
  - same for the efficiency ratio

$$\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-) = N_{d(s)} \times \frac{f_u}{f_{d(s)}} \times \frac{1}{\mathcal{D}_{\rm norm}} \times \left[\mathcal{B}(B^+ \to J/\psi K^+) \times \mathcal{B}(J/\psi \to \mu^+\mu^-)\right]$$

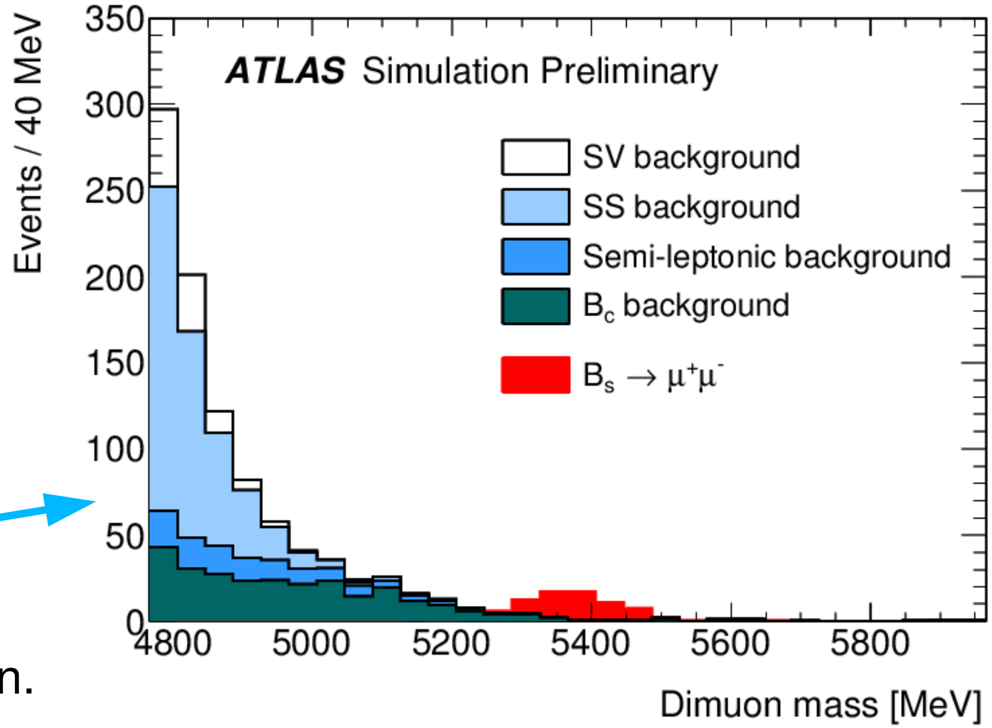
$$\mathcal{D}_{\rm norm} = \sum_k N^k_{J/\psi K^+} \alpha_k \left(\frac{\varepsilon_{\mu^+\mu^-}}{\varepsilon_{J/\psi K^+}}\right)_k$$

- index *k* runs on the trigger and data categories
- α_k takes into account the prescaling factors

Background contributions

In order of relative amplitude:

- combinatorial background from opposite-side muons:
  - dominant component
  - with smooth distribution across the dimuon invariant mass range
- partially reconstructed B decays:
  - Same Vertex (SV): $B \to \mu\mu X$ decays like $B \to K^*\mu\mu$
  - Same Side (SS): semileptonic decay cascades ($b \to c\mu\nu \to s(d)\mu\mu\nu\nu$)
  - B_c decays: like $B_c \to J/\psi\, \mu\nu$
  - all these accumulate at low values of the dimuon invariant mass
  - constituted by real muons

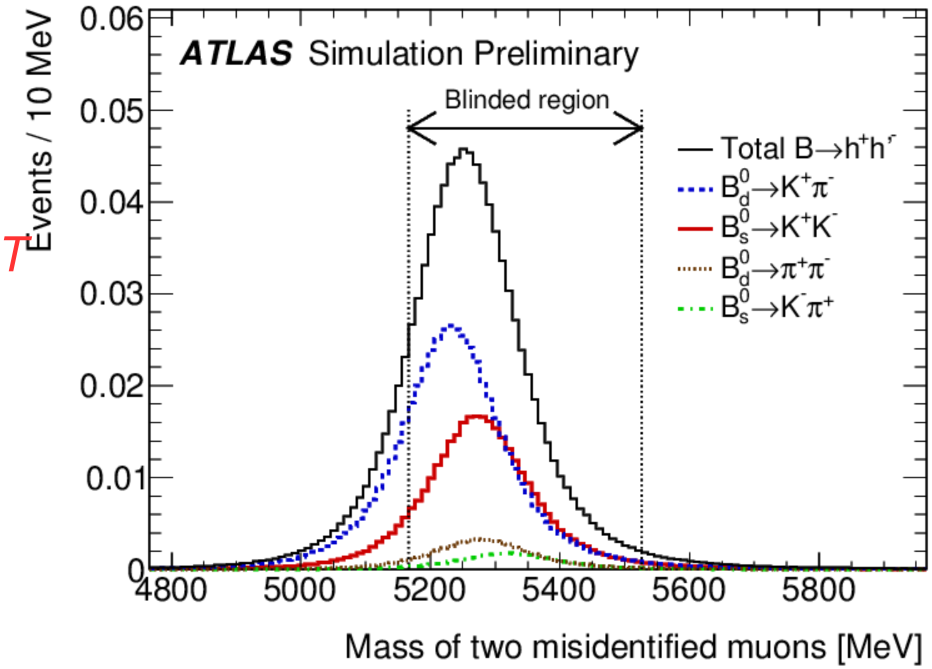


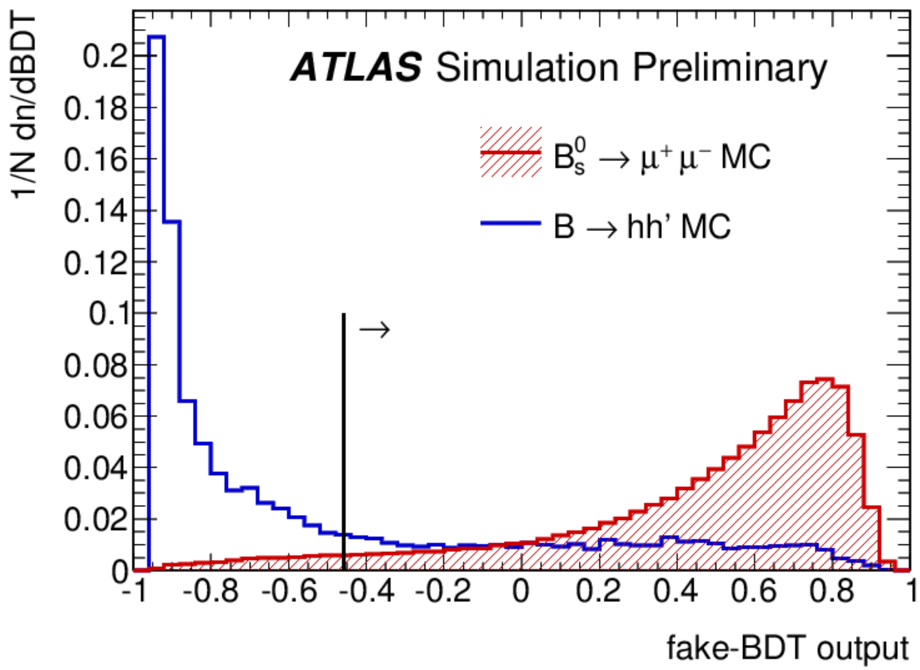
- semileptonic B and B_S decays:
  - one real muon and a charged hadron.
- peaking background from charmless hadronic $B_{(S)}$ decays:
  - B decays into two hadrons h (kaons and pions): $B^0_{(S)} \to hh'$
  - smaller component, but perfectly overlaid with the signal in dimuon invariant mass

Fake-BDT classifier against hadron misidentification

- studied on simulated samples of B →hh', signal B → μμ, and Λ_b → ph
- validated with data from φ → KK and $B^{\pm}$ → J/ψ $K^{\pm}$ decays.
- the probability of misidentification of protons is negligible (< 0.01%)
- the probability of misidentification is about 0.28% for kaons and 0.12% for pions.

reduced by a factor 0.4 with a dedicated *fake-BDT* with an efficiency of prompt muons set at 95%

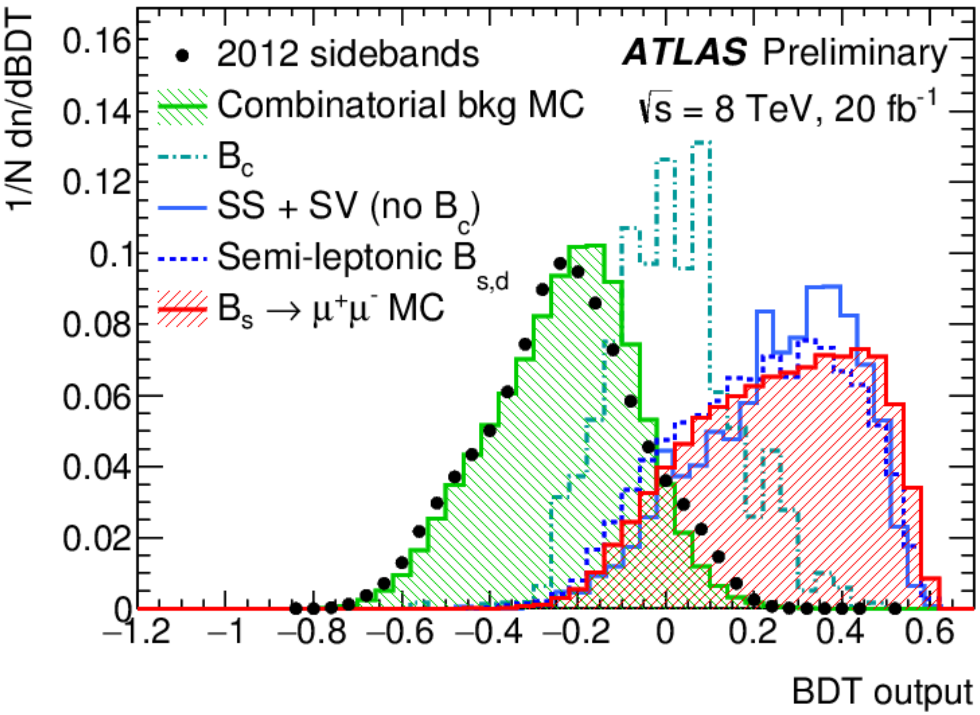




Use $B^{\pm}$ → J/ψ$K^{\pm}$ yield and efficiency ratio to normalise B →hh' (like for the signal): the total number of peaking-background events feeding into our events is 1.0±0.4

Continuum-BDT classifier against combinatorial bkg

- combinatorial background: muon pairs from uncorrelated decays of hadrons produced in the hadronisation of b and $\bar{b}$ quarks (or c and $\bar{c}$ quarks).
- separated from signal with a MVA classifier:
- 15 variables related to the B candidate, to the muons from the B candidate, to the other tracks from the same collision and to pile-up vertices.

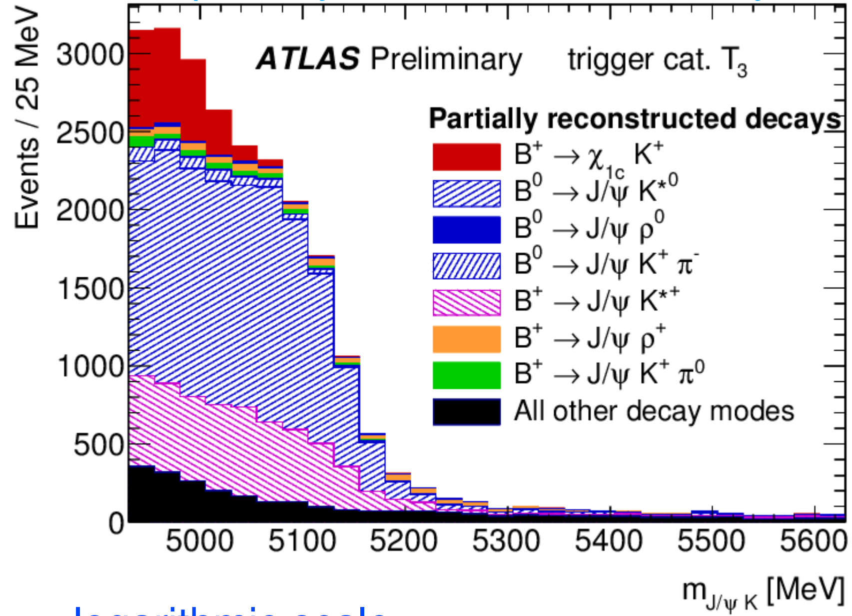


- training of the continuum-BDT done on a large MC sample of uncorrelated b (c) hadrons and $\bar{b}$ ($\bar{c}$) hadrons with forced decays into final states containing muons: 1.4G MC events
- tested on high-mass sideband data: not perfect data-MC agreement, but sample good enough for training, which is the sole use of this sample.
- B-related backgrounds behave like signal: SS-SV, semileptonic decays, peaking background

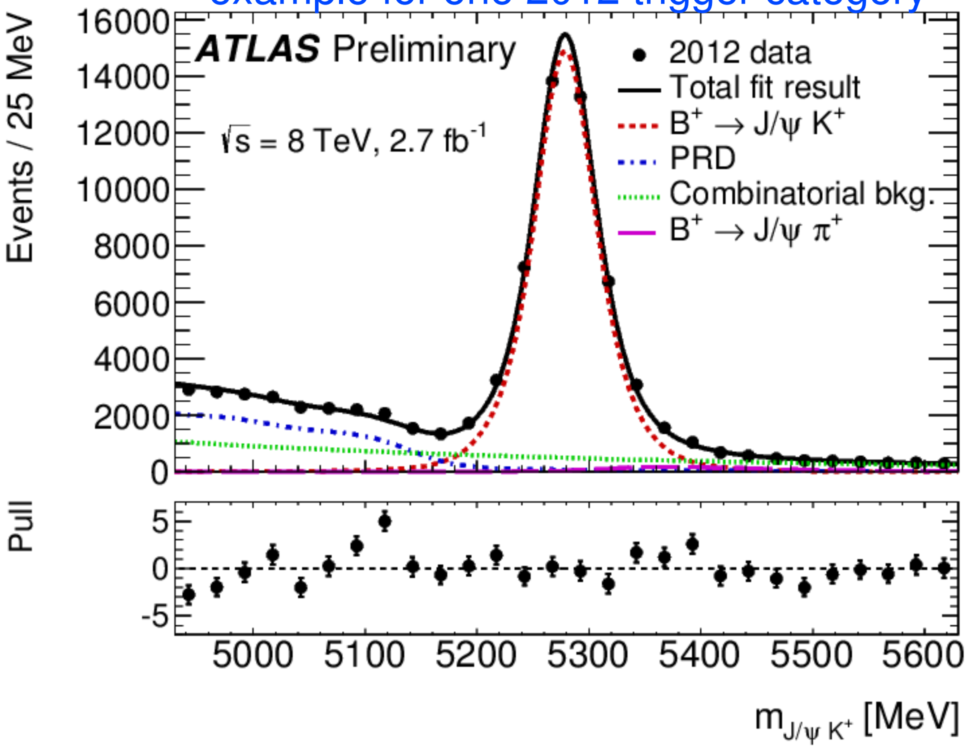
Normalisation B yield extraction

- applied fake-BDT and continuum-BDT selections (optimised for signal)
- yields extracted separately in the 4 categories: three trigger categories for 2012 and 2011 data
- unbinned maximum likelihood fit of the invariant mass distribution: $m_{J/\psi K} \rightarrow m_{\mu\mu K}$

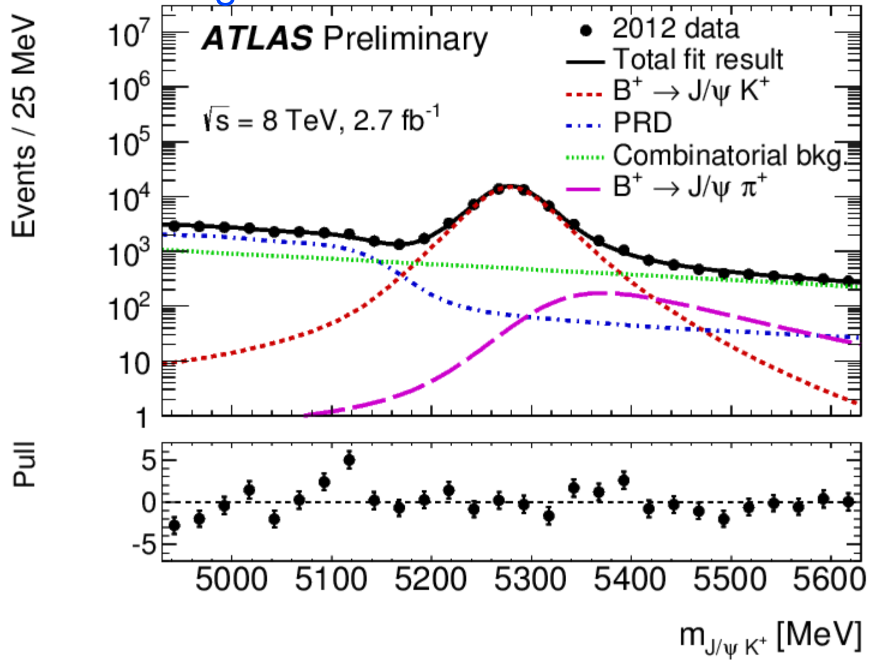
partially reconstructed B decays



example for one 2012 trigger category



logarithmic scale



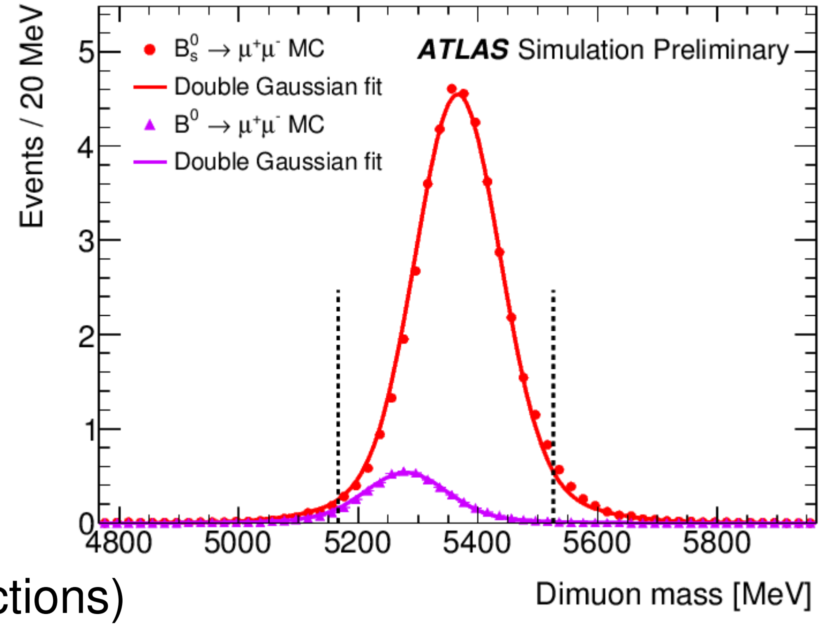
Efficiency ratio $\varepsilon_{\mu\mu}/\varepsilon_{J/\psi K}$

$$\mathcal{D}_{\mathrm{norm}} = \sum_k N^k_{J/\psi K^+} \alpha_k \left(\frac{\varepsilon_{\mu^+\mu^-}}{\varepsilon_{J/\psi K^+}} \right)_k$$

- in each category (k) the efficiency ratio is obtained from MC
- p_T and η MC spectra are tuned on data from the reference channels: included a systematic uncertainty from the statistical error from this reweighting
- residual trigger efficiencies are extracted from tag&probe studies based on J/ψ and Υ
- systematic uncertainty from data-MC discrepancies: assessed from the data-MC comparisons of the discriminating variables used in the continuum-BDT: dominant systematic contribution to D_{norm}
  - isolation (based on p_T of tracks within a cone of $\Delta R<0.7$) requires tuning in the $B^{\pm}$ mode: applied correction to the central value of the efficiency ratio.
- For B^0_S:
  - additional correction due to lifetime difference between the B^0_S mass eigenstates: lifetime taken from SM prediction and efficiency correction (+4%) taken from MC

- Total correction to the central value of the efficiency ratio: +3% for B^0 and -1% for B^0_S (including the lifetime correction)
- Total systematic uncertainty ±5.9% on the normalisation term D_{norm}

Signal yield extraction



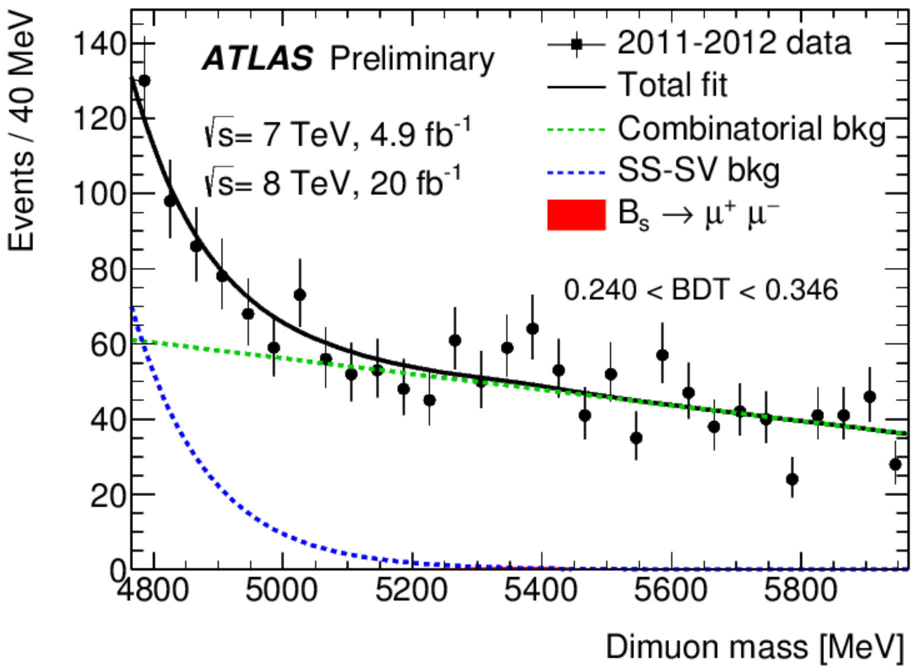
- signal yields (N_d and N_S) are extracted with a unbinned maximum likelihood fit to the dimuon invariant mass distribution
- the fit is performed simultaneously in three categories corresponding to three continuum-BDT ranges (continuum-BDT bins) chosen for constant signal efficiency (18% including corrections)

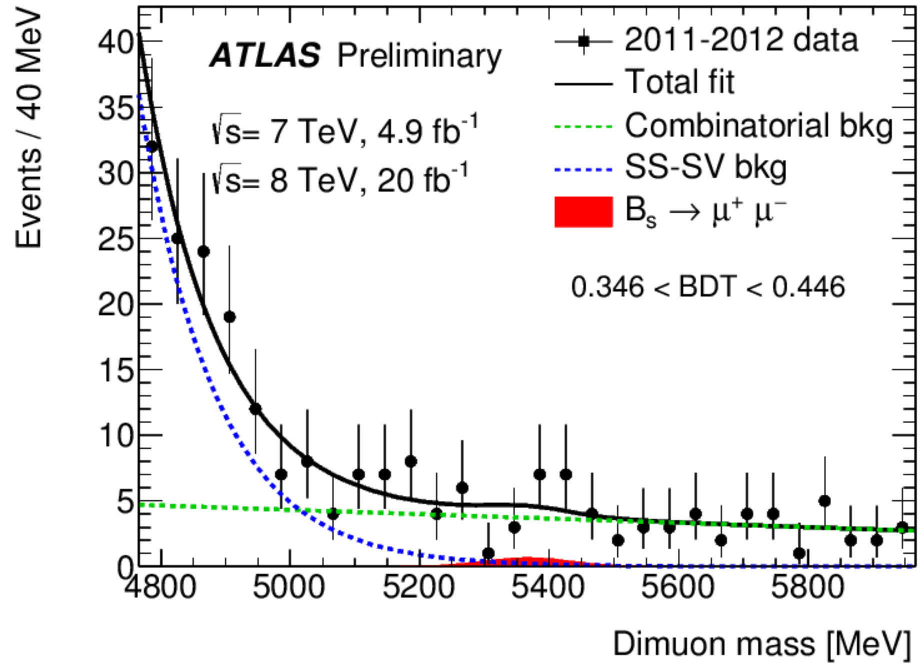
- *signal*: two Gaussian distributions with common mean, shape constrained across continuum-BDT bins and fixed to the MC shapes, varied for systematic uncertainty
- *SS-SV background*: exponential distribution, parameters floated in the fit, shape constrained across the continuum-BDT bins, independent normalisations.
- *peaking background*: two Gaussian distributions constrained across continuum-BDT bins and fixed to the MC shapes, normalisation fixed to 1.0±0.4 total events
- *continuum background*: first order polynomial, parameters floated in the fit, shape loosely constrained across the continuum-BDT bins, independent normalisations

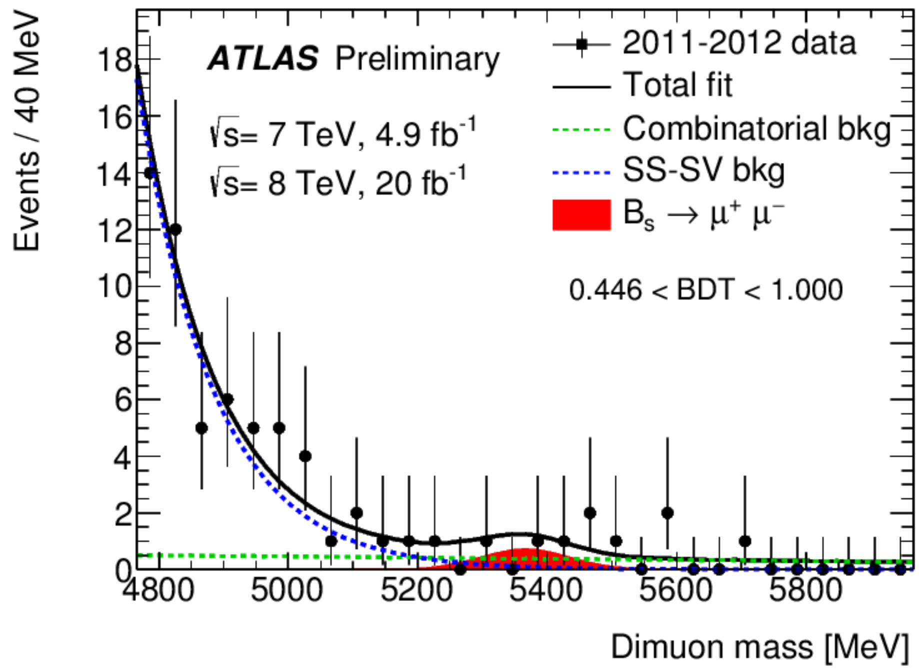
- systematics obtained by varying all the above:

$$\sigma_{syst}(N_S) = \sqrt{2^2 + (0.06 \times N_S)^2} \text{ and } \sigma_{syst}(N_d) = 3 \text{ events}$$

Signal yield extraction







- yields constrained to be positive:
  - $N_S = 11$ and $N_d = 0$
- no contraints on positive yields:
  - $N_S = 16 \pm 12$ and $N_d = -11 \pm 9$
- fewer B^0_S events than expected
- no B^0 events
- Expected signal from SM predictions:
  - $N_S = 41$ and $N_d = 5$

Branching fraction extraction

$$\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-) = N_{d(s)} \times \frac{f_u}{f_{d(s)}} \times \frac{1}{\mathcal{D}_{\text{norm}}} \times \left[\mathcal{B}(B^+ \to J/\psi K^+) \times \mathcal{B}(J/\psi \to \mu^+\mu^-)\right]$$

$$\mathcal{D}_{\text{norm}} = \sum_k N^k_{J/\psi K^+} \alpha_k \left(\frac{\varepsilon_{\mu^+\mu^-}}{\varepsilon_{J/\psi K^+}}\right)_k$$

The normalization includes:

- $B^\pm$ branching fraction (world averages)
- the fragmentation fraction f_u/f_S from the ATLAS measurement of f_S/f_d performed in the same p_T, η range: 0.240 ± 0.020 (8% systematic)
- the efficiency ratios and $B^\pm$ yields in the D_{norm} term
- The total uncertainty in the normalisation is
  - ±11% for BR($B^0_S \to \mu^+\mu^-$)
  - ± 7% for BR($B^0 \to \mu^+\mu^-$)

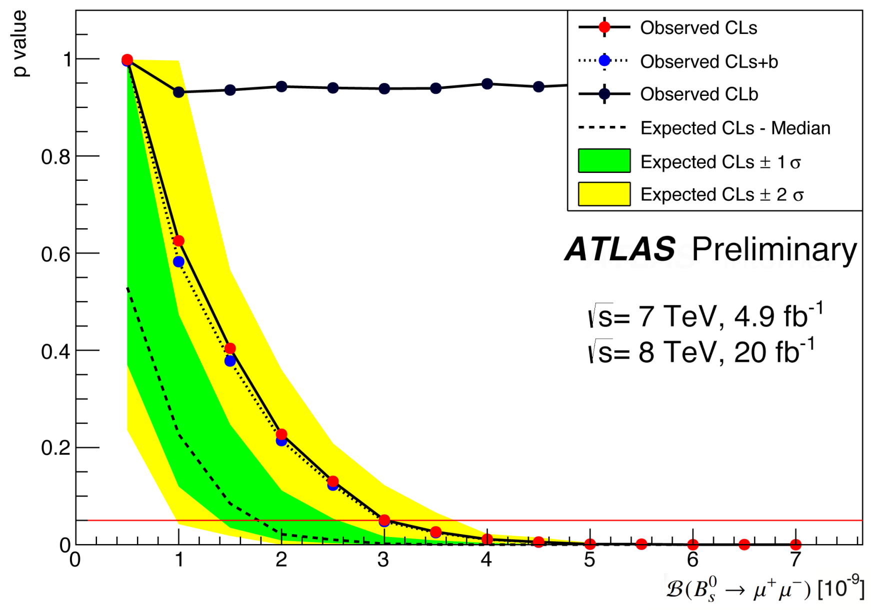
ATLAS
PRL 115 (2015) 262001

Result for the B^0_s branching fraction

- central value is obtained within the boundary of non-negative branching fractions
- the errors are obtained by means of a frequentist belt, using pseudo-MC experiments and include both statistic and systematic error. The systematic uncertainty is
  - $\sigma_{syst} = \pm\, 0.3 \times 10^{-9}$

$$\mathbf{B(B^0_s \to \mu^+\mu^-) = 0.9\,^{+1.1}_{-0.8} \times 10^{-9}}$$

- the upper limit from CLs is
  - $B(B^0_s \to \mu^+\mu^-) < 3.0 \times 10^{-9}$ at 95% CL
- the observed compatibility with the null hypothesis (only background) corresponds to
  - $p = 0.08\ (1.4\sigma)$
- the expectation for a SM signal is
  - $p = 0.0011\ (3.1\sigma)$



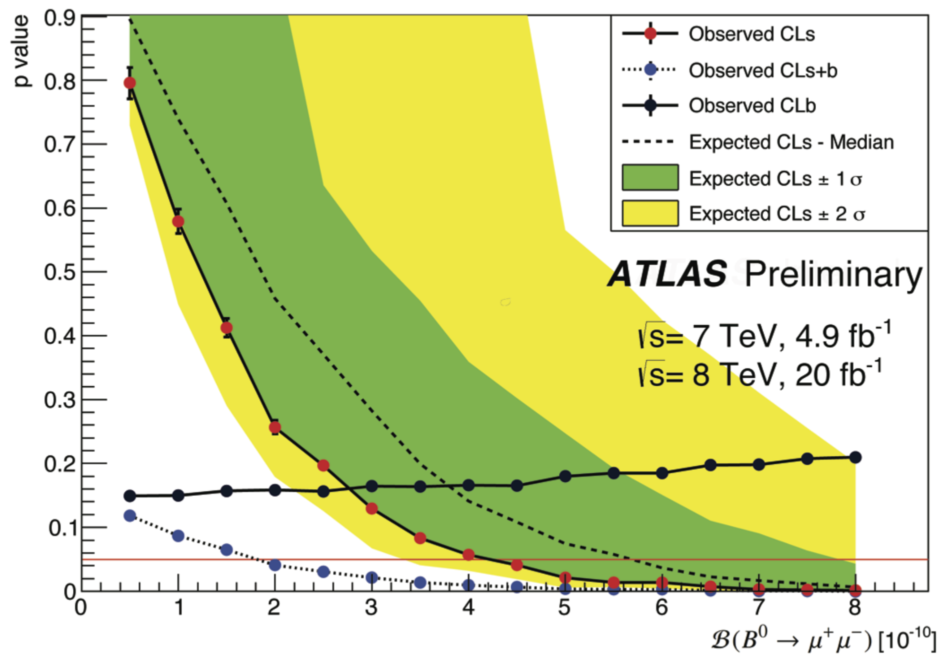
Result for the B^0 branching fraction

- upper limit set using CLs technique

$$\mathbf{B(B^0 \to \mu^+\mu^-) < 4.2 \times 10^{-10} \text{ at } 95\% \text{ CL}}$$

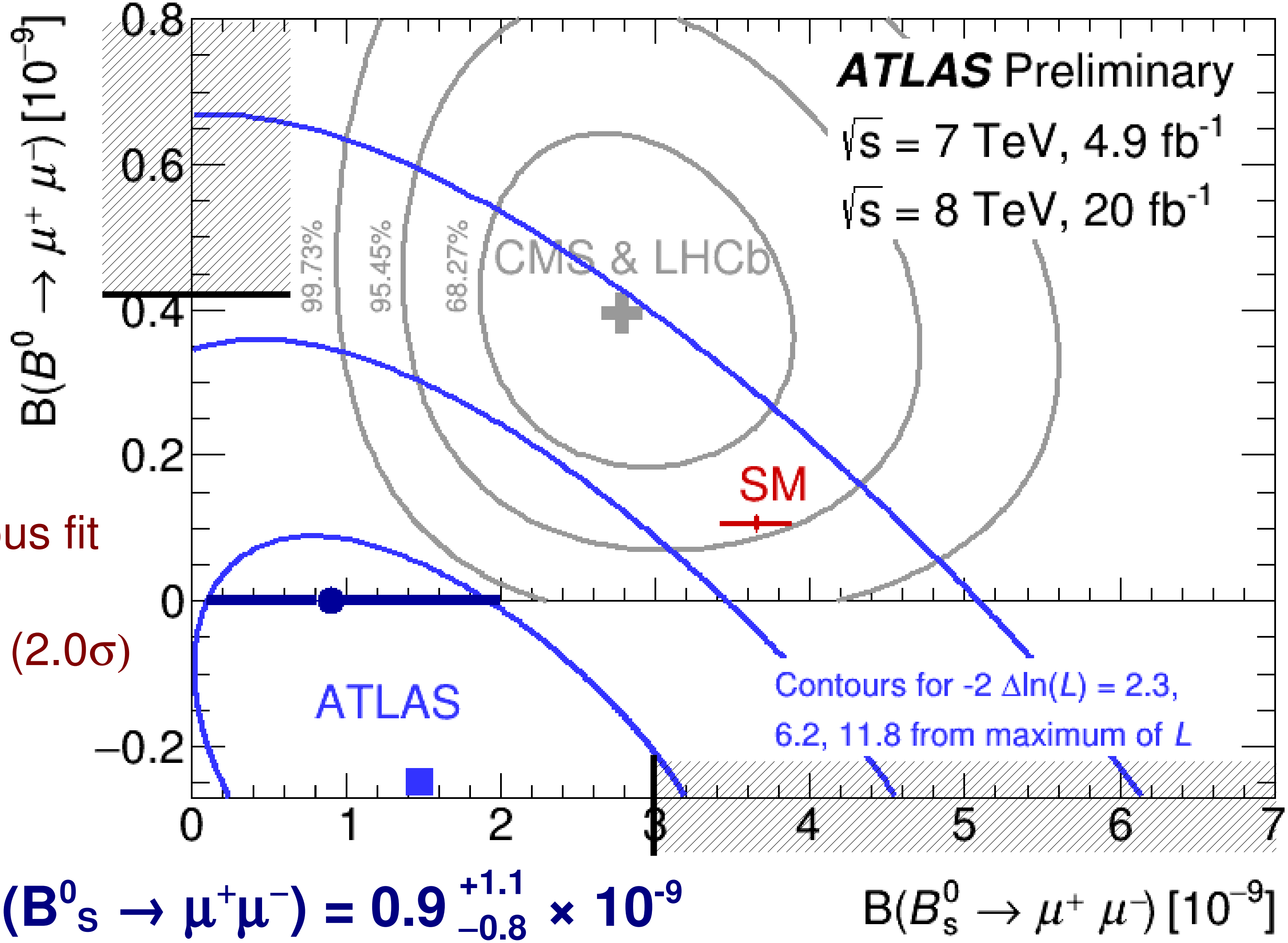
- no signal, $B(B^0_S \to \mu^+\mu^-)$ left free to be determined in the fit
- CLb is ≈ 0.15 for $B(B^0 \to \mu^+\mu^-)$ near 0:
  - -1σ fluctuation of background
- expected limit < $5.7^{+2.1}_{-1.2} \times 10^{-10}$

- the limit is higher than the SM prediction
  - $B(B^0)_{SM} = (1.06 \pm 0.09) \times 10^{-10}$
- the expected significance for $B(B^0 \to \mu^+\mu^-)$ assuming the SM branching fraction is 0.2σ



Conclusions

$\mathcal{B}(B^0 \to \mu^+\mu^-) < 4.2 \times 10^{-10}$ at 95% CL



compatibility
of the simultaneous fit
with the SM:
p-value = 0.048 (2.0σ)

$\mathcal{B}(B^0_s \to \mu^+\mu^-) = 0.9^{+1.1}_{-0.8} \times 10^{-9}$

back-up slides